

Key analysis errors and airborne wind lidar observations

ROLAND KOCH^{1,2}, MARTIN WEISSMANN^{*2} and MARTIN EHRENDORFER³

¹Institute of Meteorology and Geophysics, University of Innsbruck, Austria

²Institute of Atmospheric Physics, DLR Oberpfaffenhofen, Wessling, Germany

³Department of Meteorology, The University of Reading, United Kingdom

(Manuscript received January 28, 2007; in revised form September 24, 2007; accepted October 29, 2007)

Abstract

Inaccurate initial conditions can produce significant forecast failures of numerical weather prediction models. An iterative algorithm that uses the adjoint forecast model and is aimed at minimizing the forecast error leads to the so-called key analysis errors (KAEs). Assuming that forecast error growth is dominated by the analysis error, the KAEs are assumed to represent that part of the analysis error that is primarily responsible for a poor forecast. Thus, KAEs should indicate how to improve an analysis. In addition, analysis errors can be identified by monitoring the differences of observations and analysis fields (analysis departures). The purpose of this study is to gain a further understanding of the structure of KAEs and to investigate the question to what degree KAEs are related to analysis errors. Airborne Doppler wind lidar (DWL) observations over the Northern Atlantic collected during the Atlantic THORPEX Regional Campaign (A-TReC) are analysed to evaluate these considerations. These observations were passively monitored and actively assimilated in experiments using the ECMWF global model to form the basis for the computation of analysis departures and analysis differences. Results confirm that analysis departures and KAEs optimized for both the Northern hemisphere and a predefined forecast domain are not correlated. Qualitative comparisons also show large differences in structure and magnitude of KAEs and analysis departures. Nevertheless, primarily in view of the different magnitude of KAEs and analysis departures, there is no basis for rejecting the hypothesis that KAEs are actually embedded in, and an important fast-growing part of the true analysis errors.

Zusammenfassung

Unsicherheiten in den Anfangsbedingungen eines numerischen Wettervorhersagemodells können zu signifikanten Vorhersagefehlern führen. Ein iterativer Algorithmus, basierend auf dem adjungierten Vorhersagemodell, der zur Minimierung des Vorhersagefehlers verwendet wird, liefert sogenannte Key Analysis Errors (KAEs bzw. "Schlüssel-Analysefehler"). Unter der Annahme, dass das Anwachsen des Vorhersagefehlers durch den Analysefehler dominiert wird, stellen die KAEs jenen Teil des Analysefehlers dar, der hauptverantwortlich für den Vorhersagefehler ist. Die KAEs sollten daher Auskunft darüber geben, wie die Analyse verbessert werden kann. Der Analysefehler kann ergänzend durch Differenzbildung von Beobachtungen und Analysefeldern identifiziert werden (sog. Analyseabweichungen). Die Zielsetzung dieser Studie ist, ein besseres Verständnis der Struktur der KAEs zu erlangen und der Frage nachzugehen, inwieweit KAEs in Zusammenhang mit Analysefehlern stehen. Flugzeug-getragene Doppler-Wind-Lidarmessungen, die während der Atlantic THORPEX Regional Campaign (A-TReC) über dem Nordatlantik durchgeführt wurden, werden hier analysiert, um diese grundlegenden Überlegungen zu evaluieren. Diese Messungen wurden passiv überwacht und aktiv in Experimenten mit dem globalen ECMWF Modell assimiliert; sie bilden die Basis für die Berechnung von Analyseabweichungen und Analysedifferenzen. Die Ergebnisse bestätigen, dass Analyseabweichungen und KAEs, optimiert für die Nordhemisphäre und für ein vordefiniertes Vorhersagegebiet, nicht korrelieren. Qualitative Vergleiche zeigen auch große Unterschiede in Struktur und Größenordnung von KAEs und Analyseabweichungen. Dennoch, und zwar in erster Linie auf Grund der unterschiedlichen Größenordnung von KAEs und Analyseabweichungen, liefert die Studie keine Grundlage dafür, die Hypothese zu verwerfen, da die KAEs im wahren Analysefehler eingelagert sind und einen wichtigen, rasch anwachsenden Teil des Analysefehlers darstellen.

1 Introduction

The Atlantic THORPEX^{**} Regional Campaign (A-TReC) during late autumn 2003 aimed at testing the possibilities offered by an observational network performing additional targeted observations. Among other

observing systems, an airborne 2- μ m Doppler wind lidar (DWL) of the Deutsches Zentrum für Luft- und Raumfahrt (DLR) operated during eight research flights (WEISSMANN et al., 2005). The DWL measurements yield vertical cross-sections of wind beneath the aircraft with a horizontal and vertical resolution of 5–40 km and 100 m, respectively, depending on system configuration. They have an observational error of the zonal and meridional wind component of about 0.75–1 m/s. Recent studies show that the assimilation of such airborne DWL

^{*}Corresponding author: Martin Weissmann, Institute of Atmospheric Physics, DLR Oberpfaffenhofen, 82230 Wessling, Germany, e-mail: martin.weissmann@dlr.de

^{**}Information on the THORPEX Programme is available at <http://www.wmo.int/thorpeX>

measurements can significantly reduce the forecast error of NWP models (WEISSMANN and CARDINALI, 2007).

Such reductions of forecast error are achievable through improved analyses which emphasizes the role of the data assimilation process, as well as the importance of the availability of accurate observations. Improving analyses implies at the same time a reduction of the analysis error (AE). Little is known about the “true” AE and approximations are necessary in order to investigate its nature (see, for example, ERRICO et al., 2007). In the present study, the above-mentioned DWL measurements are used to evaluate an approximation to the AE, and to subsequently compare that approximation with the so-called key analysis errors (ISAKSEN et al., 2005) that are sometimes taken as an adjoint-based measure of the AE.

A comprehensive approach to the investigation of AEs is available through adjoint-based sensitivity studies (ERRICO, 1997; JUNG and BARKMEIJER, 2006; ISAKSEN et al., 2005). This approach allows quantitative assessment of the sensitivity of a numerical weather prediction system to the initial conditions. Since uncertainty in the initial conditions is a main source of unpredictability, it is implied that small changes in the initial conditions may produce significant changes in the final model forecast.

The response of a given NWP model to observations depends on the observing system and the data assimilation method used. The information available for estimating the initial state is generally insufficient due to incomplete observations in data sparse areas and observational errors. This is partly compensated for through the data assimilation process that combines a short-term forecast with observations. Real data assimilation systems, however, are always imperfect and may introduce errors as well. Furthermore, the use of an imperfect NWP model introduces model-related errors.

Adjoint techniques allow for quantitatively identifying the error structures in the initial condition to which the forecast error is (most) sensitive. In this approach, an objective cost function measuring the misfit between a short-term forecast and the verifying analysis is minimized to determine the structures to which the forecast errors are most sensitive. The backward integration of the adjoint of the tangent-linear version of the nonlinear model gives the sensitivity gradient.

The gradient obtained in this way can then be applied to an iterative algorithm, which results in small perturbations minimizing the cost function. These initial-time perturbations are known as key analysis errors (KAEs) and they can be used to update the analysis in an effort to obtain a more accurate forecast. As such, KAEs are supposed to be the fast-growing components of the AE that, when time-evolved, dominate the forecast error (KLINKER et al., 1998). Previous studies, such as those

carried out at the European Centre for Medium-Range Weather Forecasts (ECMWF), have shown that adding KAEs to the analysis reduces the forecast error, but at the same time draws the analysis further away from observations assimilated in the operational ECMWF analysis (ISAKSEN et al., 2005).

Little is known about the qualitative and quantitative similarities of KAEs and AEs. In order to provide more evidence on the interpretation and nature of KAEs, the present study investigates the correlation between KAEs and AEs, evaluated on the basis of the DWL measurements referred to above. As the DWL observations were shown to have a statistical error of only 0.75–1 m/s and no relevant bias (WEISSMANN et al., 2005), the differences between DWL observations and the analysis (observation-analysis departures) may be taken as a reasonably good estimate of the AE. The present comparison includes a statistical approach and also the comparison of structural differences on vertical cross-sections measured by the airborne DWL.

Given such an estimate of the AEs, the present paper investigates the hypothesis that KAEs are consistent with present knowledge about the AEs. Or, formulated slightly differently, the question is asked whether there is evidence for rejecting *the hypothesis that the KAEs are embedded in the AEs and thus are an important part of AE*.

It has been demonstrated in various contexts (e.g., GELARO et al., 2000, and GELARO et al., 2002) that a very small (in terms of magnitude) portion of (approximations to) AE grows into significant portions of the forecast error (see, e.g., Fig. 11 of GELARO et al., 2000). In addition, these small portions are found to be – in terms of their structures – significantly different from the much larger non-growing part of the AE. As such, “key” growing structures of the AE have been found to be embedded in the true AE and may thus be hidden beneath a much larger non-growing part of the AE. Such behaviour of growing structures is indeed very consistent with the nonmodal nature of error growth, when investigated in the presence of complex and time-varying basic states.

On the basis of these results, poor correlations found between AEs and KAEs, or the fact that KAEs may not have “realistic” structures (CARON et al., 2007) does not provide a basis for rejecting the above hypothesis. Still, however, such findings emphasize at the same time concerns that AEs and KAEs are completely different quantities, which would strongly question their use in Sensitivity Observing System Experiments (SOSEs, see, MARSEILLE et al., submitted).

Through their construction which is also based on a rather low-resolution model version, KAEs are expected to mainly represent large-scale variations, whereas observation departures are a combination of various scales,

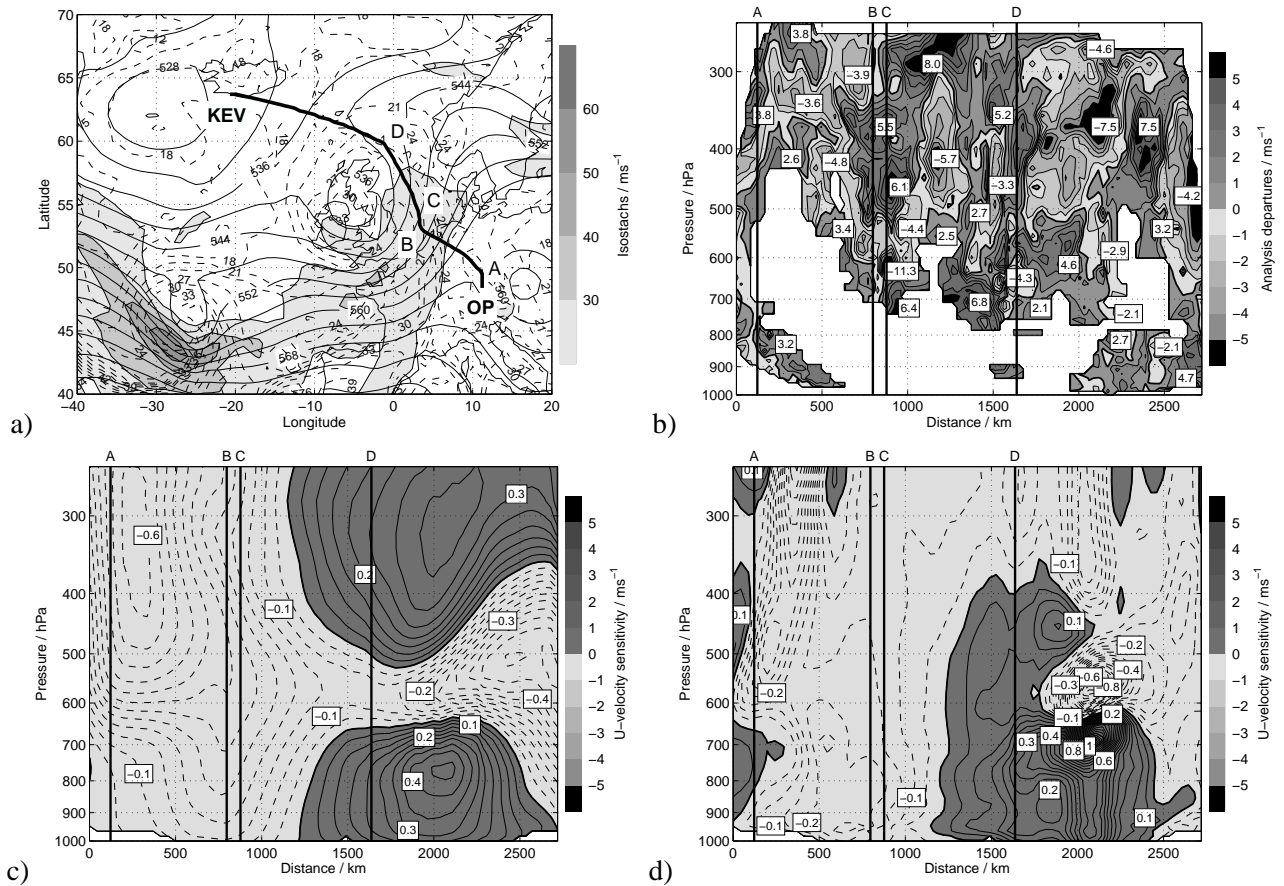


Figure 1: Flight track of the DLR Falcon from Oberpfaffenhofen (OP), Germany to Keflavik (KEV), Iceland. Operational ECMWF analyses 1200 UTC 14 November 2003 of geopotential height (10m) at 500 hPa (denoted by solid contours drawn at an interval of 40 m), isostachs (ms^{-1}) at 500 hPa (denoted by shaded contours drawn at an interval of 10 ms^{-1}) and equivalent potential temperature ($^{\circ}\text{C}$) at 850 hPa (denoted by dashed contours drawn at an interval of 3 $^{\circ}\text{C}$) are overlaid. The turning points of the DLR Falcon are indicated by A, B, C and D (a). Vertical cross-sections of u -velocity analysis departures (ms^{-1}) along the flight track (b), B-norm KAEs (ms^{-1}) (c) and TE-norm KAEs (ms^{-1}) (d). For clarity of presentation, local minima and maxima are shown in panel (b). KAEs are bilinearly interpolated to the flight track. Shaded contour interval in panels (b), (c) and (d) is 1 ms^{-1} . Additional contour interval in panels (c) and (d) is 0.05 ms^{-1} , with positive values shown as solid contours and negative values shown as dashed contours. The zero contours are drawn by thick black lines in panels (b), (c) and (d). Vertical black lines indicate turning points of DLR Falcon at location A, B, C and D. In the white areas of panel (b) no lidar data were available.

often dominated by smaller scales (as should be expected in case of uncorrelated observation errors). Thus, in the present study, KAEs are also compared to differences of analyses obtained from two assimilation experiments, one of which uses lidar data, whereas the other does not. These analysis differences mainly contain the large-scale variations of the observation departures.

Several uncertainties arise, when observation departures and analysis differences are compared to KAEs. These include interpolation errors, data-representativeness and observational errors (although the lidar data is the most accurate wind information available as shown by WEISSMANN and CARDINALI, 2007). As a consequence, no direct alignment between KAEs and observation departures can be expected. Instead, the study focuses on qualitative differences and similarities of their structures. Furthermore, their

spectra are compared. Before discussing the observational data set and the comparison of these observations with KAEs in sections 3–5, a brief clarification of methodology and notation relevant for KAEs is given.

2 Quantifying analysis errors

2.1 Linear error growth

The prognostic equation of an NWP model has typically the form:

$$\frac{d\mathbf{x}}{dt} = \mathcal{M}(\mathbf{x}), \quad (2.1)$$

where \mathbf{x} is the model state vector, and \mathcal{M} is the non-linear forecast model operator, respectively. The time evolution of initial perturbations $\delta\mathbf{x}_0$ along the nonlinear trajectory $\mathbf{x}(t)$ is to first-order accuracy described by

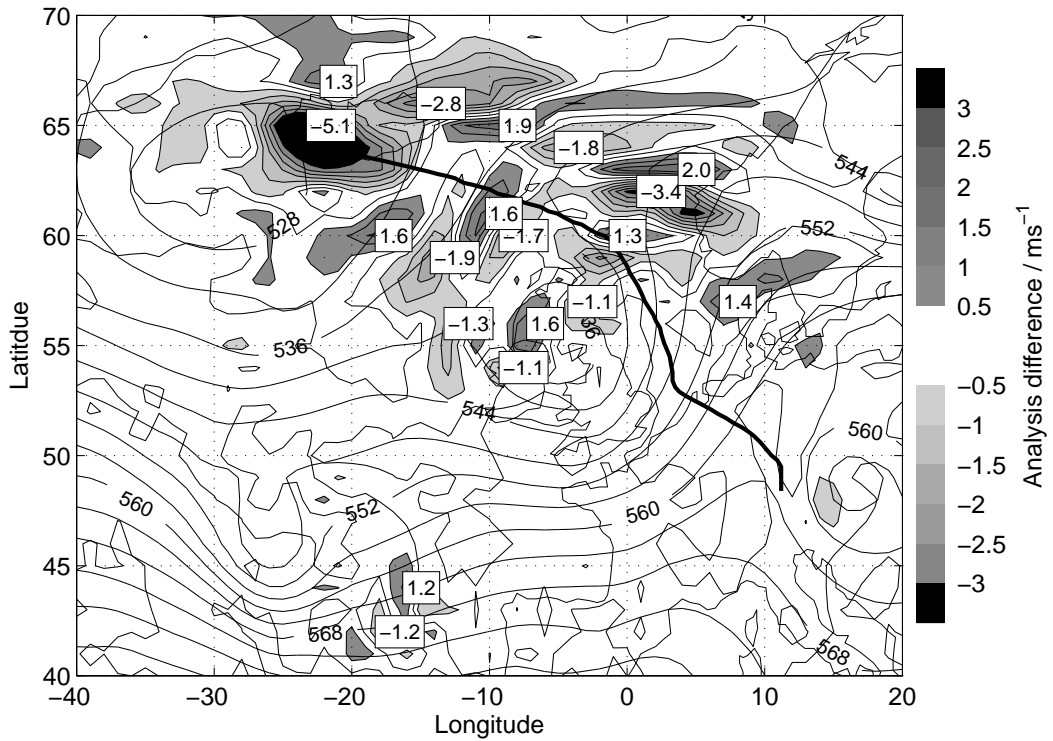


Figure 2: ECMWF 1200 UTC 14 November 2003 u -velocity field of analysis difference (ms^{-1}) at model level 39 (approx. 500 hPa). Operational ECMWF analysis of geopotential height (10m) at 500 hPa (denoted by solid contours drawn at an interval of 40 m) is overlaid. Shaded contour interval is 0.5 ms^{-1} . The zero contour is suppressed. For clarity of presentation, local minima and maxima are shown. Thick black line indicates the flight track.

the tangent linear equation corresponding to Eq. (2.1), given by:

$$\frac{\partial \delta \mathbf{x}}{\partial t} = \mathbf{L} \delta \mathbf{x}, \quad (2.2)$$

where \mathbf{L} is the Jacobian of \mathcal{M} , that represents the differentiations of \mathcal{M} with respect to \mathbf{x} and is known as the tangent linear operator of the nonlinear model. The solution of Eq. (2.2) can be written as:

$$\delta \mathbf{x}_t = \mathbf{M}(t, t_0) \delta \mathbf{x}_0, \quad (2.3)$$

where $\mathbf{M}(t, t_0)$ represents the forward tangent linear propagator (resolvent), which evolves the initial perturbation $\delta \mathbf{x}_0$ along the nonlinear trajectory $\mathbf{x}(t)$ from the initial time t_0 to the optimization time t (final time of forecast). The propagator $\mathbf{M}(t, t_0)$ depends on the initial condition (analysis) \mathbf{x}_0 . Thus, Eq. (2.3) describes the linear growth of small errors present in the initial condition.

2.2 Observation departures

The ECMWF model system uses a four-dimensional variational (4D-Var) data assimilation system. The objective of 4D-Var is to iteratively minimize a quadratic cost function that measures the degree of fit of a model trajectory (short-term forecast) to a background state and to observations over the assimilation window (currently 12 hours). The full horizontal resolution (spectral triangular truncation at wavenumber 511 denoted as

T511) and vertical resolution (sixty levels denoted as L60) of the ECMWF model system are used for comparing observation and model equivalents.

In the minimization of the cost function (RABIER et al., 2000) through the incremental formulation (COURTIER et al., 1994) the resolution is T95/T159. In the iterative minimization problem, the vector of *background departures* (innovation vector) is obtained in the first full non-linear model integration using the background state \mathbf{x}_b , that summarises all information used before as initial condition and has the form:

$$\mathbf{d}_i = \mathbf{y}_i - \mathcal{H}_i(\mathbf{x}_b), \quad (2.4)$$

where i represents the index of observation time, \mathbf{y}_i is the observation vector and \mathcal{H}_i is the nonlinear observation operator, that maps model grid point values to observation space at time index i and includes the model integrations.

The *analysis departures* (analysis residuals) are obtained by replacing the background state \mathbf{x}_b with the analysis \mathbf{x}_a in Eq. (2.4) as:

$$\mathbf{d}_i = \mathbf{y}_i - \mathcal{H}_i(\mathbf{x}_a). \quad (2.5)$$

The analysis departures are obtained in the final nonlinear model integration using the optimal solution to the minimization problem as initial condition. They measure how well the analysis fits the observations.

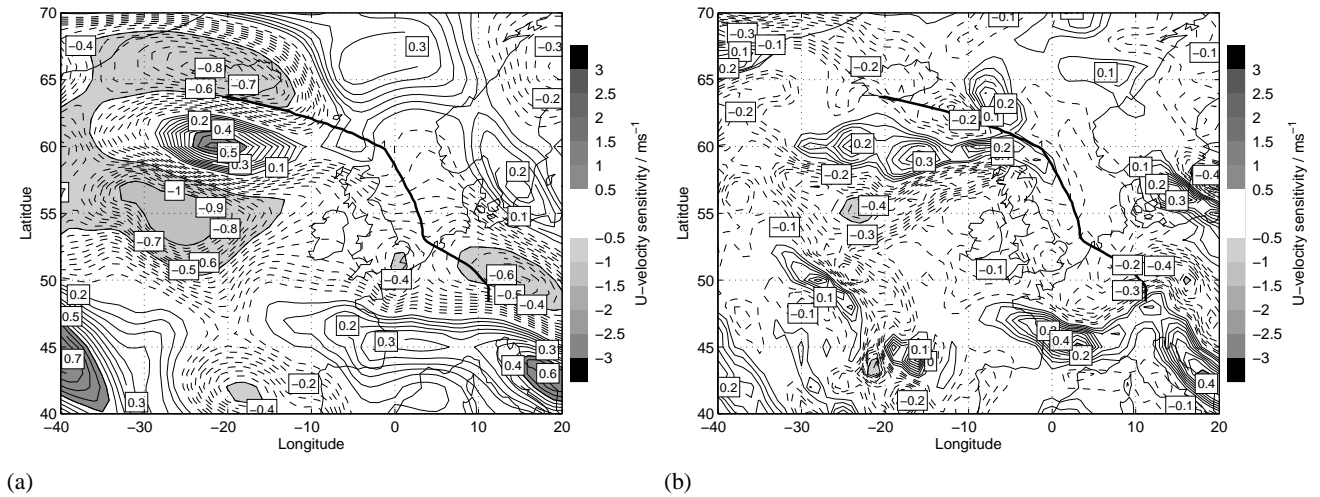


Figure 3: Sensitivity fields 1200 UTC 14 November 2003 at model level 39 (approx. 500 hPa) of u -velocity B-norm KAEs (ms^{-1}) (a) and TE-norm KAEs (ms^{-1}) (b). Shaded contour interval is 0.5 ms^{-1} . Additional contour interval is 0.05 ms^{-1} , with positive values shown as solid contours and negative values shown as dashed contours. The zero contour is suppressed.

Passing information contained in observations, from one 4D-Var cycle to the next, is known as active assimilation, whereas passively monitored observations are detected only without minimizing the corresponding innovations.

Analysis departures of passively monitored extra observations provide a measure of the AE, representing unresolved observed variances and statistical AE. However, one must take into account, that cycling 4D-Var with additionally observed data does not guarantee an improved analysis. On average, the use of additional observations improves the analysis, but in individual events the results can be different as the assimilation system solves a statistical problem, where the errors of both the background field and the observations are unknown.

Besides the computation of analysis departures, the investigation of analysis impact of additionally observed data may provide useful information about analysis errors. A measure of the analysis impact is the computation of *analysis differences*. In this context, the analysis difference is defined as the difference between a nonlinear model run that actively assimilates additional observations and a control run with only operational observations (WEISSMANN and CARDINALI, 2007).

2.3 Key analysis errors

KAEs are defined as initial time perturbations $\delta \mathbf{x}_0$ that address the following problem:

$$\min_{\delta \mathbf{x}_0 \in \mathbb{R}^n} J(\delta \mathbf{x}_0), \quad (2.6)$$

where the cost function J is defined as:

$$\begin{aligned} J(\delta \mathbf{x}_0) &= \frac{1}{2} (\mathbf{x}_{t_f} + M\delta \mathbf{x}_0 - \mathbf{x}_{t_a})^T \mathbf{P}^T \mathbf{A} \mathbf{P} (\mathbf{x}_{t_f} + M\delta \mathbf{x}_0 - \mathbf{x}_{t_a}) \\ &= \frac{1}{2} \|\mathbf{P}(\mathbf{x}_{t_f} + M\delta \mathbf{x}_0 - \mathbf{x}_{t_a})\|_A^2, \end{aligned} \quad (2.7)$$

and the following constraint has to be satisfied:

$$\delta \mathbf{x}_0^T \mathbf{C} \delta \mathbf{x}_0 = 1. \quad (2.8)$$

The operators \mathbf{C} and \mathbf{A} are positive definite and symmetric matrices defining a norm at initial and final time; \mathbf{P} is the local projection operator targeting a specific domain (BUIZZA, 1994); \mathbf{x}_{t_f} is the result of the 48-hour integration with the full non-linear model and \mathbf{x}_{t_a} is the verifying analysis valid at final time t . A norm widely used to measure the forecast error at final time is the total energy norm (TE-norm; see ERRICO, 2000 and ISAKSEN et al., 2005).

The quadratic cost function $J(\delta \mathbf{x}_0)$ can be iteratively minimized by applying a quasi-Newton procedure. This method requires the knowledge of the gradient of the cost function with respect to $\delta \mathbf{x}_0$, which is known as the sensitivity and can be computed by using the adjoint version of the tangent linear model. A more detailed description of the sensitivity computation and of the KAEs algorithm can be found in KLINKER et al. (1998), RABIER et al. (1996) and ISAKSEN et al. (2005). Following KLINKER et al. (1998), three iterations of the minimization problem are required to achieve the final analysis perturbations, i.e., the KAEs, which improve the two-day forecast skill and best fit the observations, because KAEs are assumed to represent fastest growing components of the analysis error (see also, EHRENDORFER and

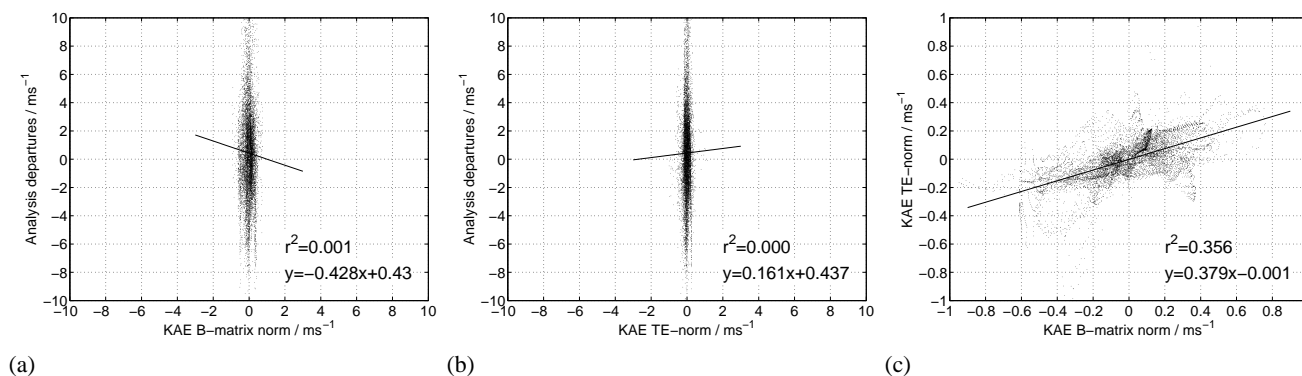


Figure 4: Correlation of u -velocity analysis departures based on passively monitored DWL data and B-norm KAEs (a) and TE-norm KAEs (b). Correlation of B-norm KAEs and TE-norm KAEs (c). The correlation sample size is 17226. The square of the sample correlation coefficient and the fitting curve are furthermore shown.

TRIBBIA, 1997 for the investigation of initial time perturbations constructed from singular vectors of the tangent linear model).

The two-day forecast skill improves with an increasing number of iterations performed within the KAEs algorithm, but the patterns obtained become more and more unrealistic. This behaviour has been investigated in more detail than given by KLINKER et al. (1998) in the study of ISAKSEN et al. (2005). In addition, KAEs are known to be very small in magnitude and it has been found that the sensitivity structures obtained depended strongly on the initial time norm used (KLINKER et al., 1998; ISAKSEN et al., 2005).

ISAKSEN et al. (2005) have investigated KAEs based on three different initial time norms induced by C. It turned out that especially during the first approximately 12 hours the sensitivity forecasts based on these three norms were further away from the observations compared to forecasts started from unperturbed initial conditions. They concluded that KAEs are not correlated to true AEs.

3 Data collection

Between 14 and 28 November 2003, the DLR Falcon research airplane made eight flights over the Northern Atlantic (WEISSMANN et al., 2005). The measured high-quality vertical wind profiles were actively assimilated and passively monitored in several experiments at ECMWF. The global model resolution for these experiments was T511/T159 (40 km/125 km) with 60 vertical levels. Analysis departures of passively monitored observations provide an independent evaluation of AEs since the forecast/analysis cycle of the model over the assimilation window is not influenced by the DWL measurements. Under the assumption that the AE over the

Northern Atlantic is relatively large due to sparse observations, extra DWL measurements are expected to contribute to an average reduction of the AE.

At ECMWF two sets of KAEs were computed covering the period 14 November through 28 November 2003 valid at 1200 UTC and 0000 UTC. The computation was performed over an optimization time interval of 48 hours using a moist version of the tangent linear and adjoint model with simplified physics. The integrations were made with a horizontal resolution of T159 (125 km) and with 60 model levels in the vertical. The final KAEs were obtained after performing six iterations of the iterative minimization algorithm. The target region was chosen to focus on the Northern Hemisphere north of 30°N. In order to illustrate the norm dependence, the TE-norm and the background-error covariance matrix norm (B-norm) were used as the initial time norms in separate computations. In both cases the constraining final time norm was the TE-norm.

Additional KAEs were computed for 1800 UTC 15 November 2003. An optimization time of 42 hours was used and the local projection operator was employed to target a predefined verification area of 2000 km by 2000 km located in the southern part of Europe. The TE-norm was used as initial and final time norm. Only three iterations were used for the iterative minimization of the cost function.

In the following, we present two selected case studies comparing magnitudes and spatial structure of KAEs with those of analysis departures based on passively monitored DWL data and analysis differences based on actively assimilated DWL data.

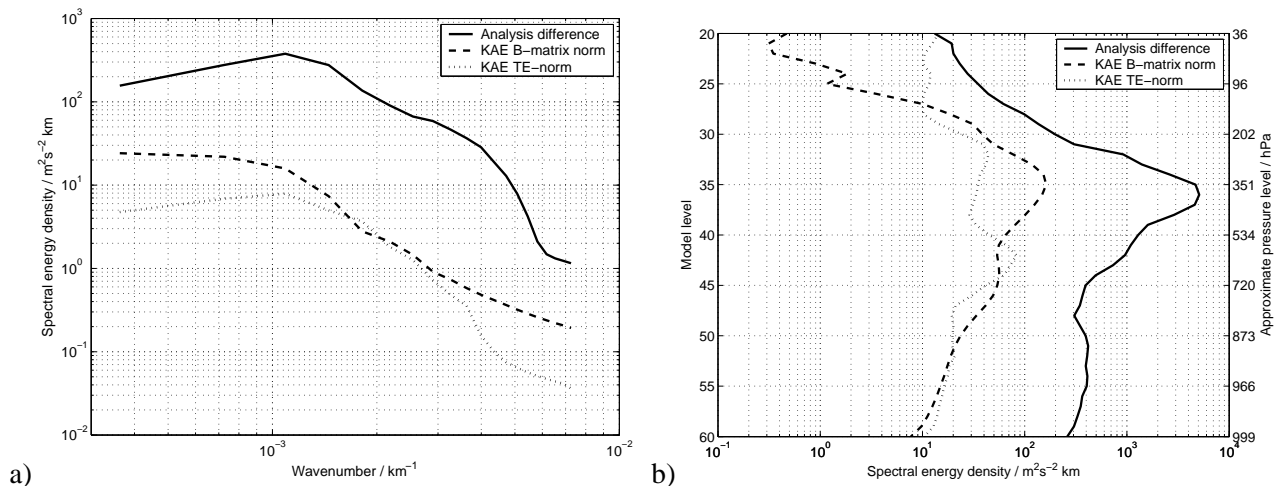


Figure 5: Horizontal mean spectral energy densities at model level 39 (approximately 500 hPa) (a) and vertical mean spectral energy density profiles (b) of analysis difference (solid line), B-norm (dashed line) and TE-norm KAEs (dotted line) for the u -velocity field. The spectra are calculated for the domain $70^\circ\text{N}/30^\circ\text{W}$ – $50^\circ\text{N}/10^\circ\text{E}$ and are valid at 1200 UTC 14 November 2003.

4 Case study for 14 November 2003

4.1 Analysis departures versus KAEs

On 14 November 2003 an upper-level trough with two embedded short-wave troughs to the west and east was situated over the Northern Atlantic. A well-developed jet stream branch on the rear side of the long-wave trough was observed as shown in Fig. 1a. A frontal system associated with the short-wave trough to the east extended from Western Europe to the North Sea. Furthermore, an area of weak low pressure was located over Central Europe. In the following days a cut-off low formed as consequence of the deepening of the western short-wave trough, drifting slowly eastward with associated cloud cover and precipitation in the south-western parts of Europe and North Africa.

The DLR Falcon research airplane performed a transfer flight from Oberpfaffenhofen, Germany to Keflavik, Iceland between 0724 and 1053 UTC on 14 November 2003 (see Fig. 1a). Figs. 1b, 1c and 1d show cross-sections of analysis departures based on passively monitored DWL measurements and corresponding KAEs for the u -velocity along the flight track of the DLR Falcon. A bilinear interpolation scheme was used to interpolate KAEs to the flight track. The u -velocity profiles observed by the DWL system have a horizontal resolution of approximately 40 km, which is similar to the horizontal resolution of the global non-linear T511/159L60 version of the ECMWF forecast model. The time difference between the two quantities is assumed to be negligibly small due to the slowly varying characteristics of large-scale flow.

The analysis departures show many fine-scale details (Fig. 1b). White gaps in Fig. 1b represent areas, where no measurements were possible due to opaque cloud cover. Comparison of the pattern of analysis departures

with corresponding KAEs based on the B-norm (Fig. 1c) and TE-norm (Fig. 1d) as initial time norm reveals differences in terms of spatial structure and magnitude. The horizontal and vertical structures of DWL analysis departures are characterized by locally organized minima and maxima with a peak spread between -11.3 m/s and 8 m/s. In contrast, KAEs are much smaller and of larger scale. The analysis departures are on average one order of magnitude larger than the KAEs.

One would expect that these discrepancies arise from the coarser resolution in the computational algorithm of KAEs. However, large differences are also seen after smoothing the analysis departures to a coarser resolution (not shown). Some of these differences in scale and magnitude are associated with violations of the assumptions underlying the adjoint-based algorithm used for computing the KAEs.

Similar differences are found to occur for the comparison of the v -velocity, other flight tracks and different processing of the DWL data (not shown). As stated in section 2.3, KAEs are initial time perturbations minimizing an objective cost function. This minimization requires the knowledge of the gradient of the cost function. Perhaps it is necessary to question the assumptions behind the calculation of the KAEs. For example, if the perfect-model assumption does not hold, and thus forecast error is dominated by modeling deficiencies, then interpreting KAEs as initial-condition errors would clearly not be justified.

Relatively large differences in spatial structure are apparent between B-norm and TE-norm KAEs (Figs. 1c and 1d). Similar to the findings of ISAKSEN et al. (2005), this result indicates that the features obtained strongly depend on the norm used. However, the KAEs are approximately of the same magnitude for both norms. Furthermore, the large variance of TE-norm perturbations

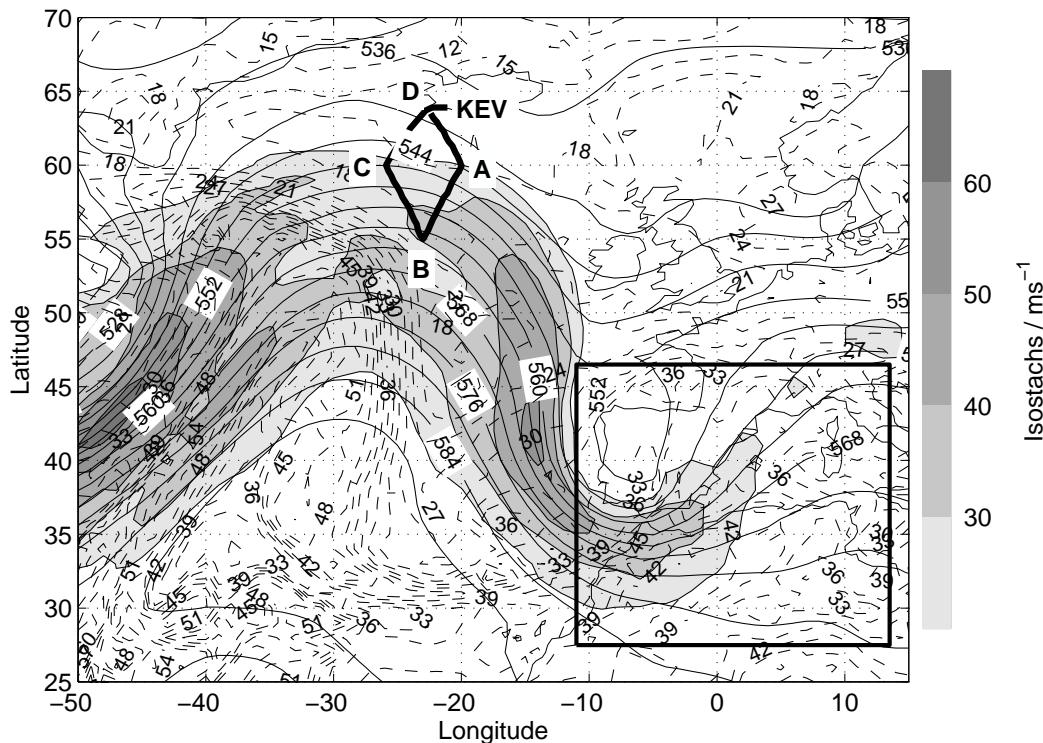


Figure 6: Same as Fig. 1a, but with flight track of the DLR Falcon on 15 November 2003. Operational ECMWF analyses are valid at 1800 UTC 15 November 2003. The black box indicates the verification area for which TE-norm KAEs are computed.

(Fig. 1d) is mainly limited to regions of the lower and middle troposphere. In comparison, the spatial structure of B-norm KAEs (Fig. 1c) is much broader and the sensitivity values are more uniformly distributed, which is mainly due to the horizontal correlation length scales of the modelled static B-matrix.

In the 4D-Var assimilation scheme, B is still prescribed statically, but implicitly acquires some flow-dependence as the model dynamics are used to evolve fields over the assimilation window. In data sparse regions (i.e., the Northern Atlantic) the information of measurements is needed to spread out to surrounding points and levels in terms of consistency, dynamics, and balance, whereas in data-dense areas, correlations in the B-matrix govern the smoothing of observed information (COURTIER et al., 1998). Therefore, B spreads out the information leading to broader and smoother patterns. The differences in the structure of B-norm and TE-norm KAEs raise the question, whether either of these norms is suitable for representing AEs.

However, KAEs are dynamical quantities and the patterns obtained are related to the synoptic situation. The vertical dipole structure of large sensitivity values around 600 hPa at approximately 2000 km away from the departing point of the DLR Falcon is probably related to the occluded frontal system located over the North Sea that is indicated by the equivalent potential temperature field (Figs. 1a, 1c and 1d). The region becomes unstable due to the moderate temperature gradi-

ent. Thus, the collocation of the KAE dipole and the occluded frontal system indicates that KAEs are linked to baroclinic instabilities (KLINKER et al., 1998; ISAKSEN et al., 2005).

At higher levels the large amplitude of B-norm KAEs probably originates from an observed cold air intrusion. The investigation of ozone mass-mixing ratio (not shown) at upper model levels indicates that ozone-rich stratospheric air was mixed into lower parts of the atmosphere. The exchange of these air masses and the involved destabilization is presumably responsible for the shape of B-norm KAEs at these levels. However, this maximum is absent in the corresponding TE-norm KAEs indicating that the shape of TE-norm perturbations is governed by low-level temperature anomalies.

In contrast, analysis departures show large values at these levels (Fig. 1b), but the extrema are organized in a much different form. The significant discrepancies in magnitude and spatial structure imply that analysis departures and KAEs represent different parts of the analysis error. The findings raise the question if KAEs represent any form of analysis errors.

4.2 Analysis impact of lidar data

The assimilation of additional observations can improve the forecast skill of NWP models, in particular if the observations are deployed in sensitive areas (LEUTBECHER, 2003, WEISSMANN and CARDINALI, 2007). The basic idea is to improve initial conditions and

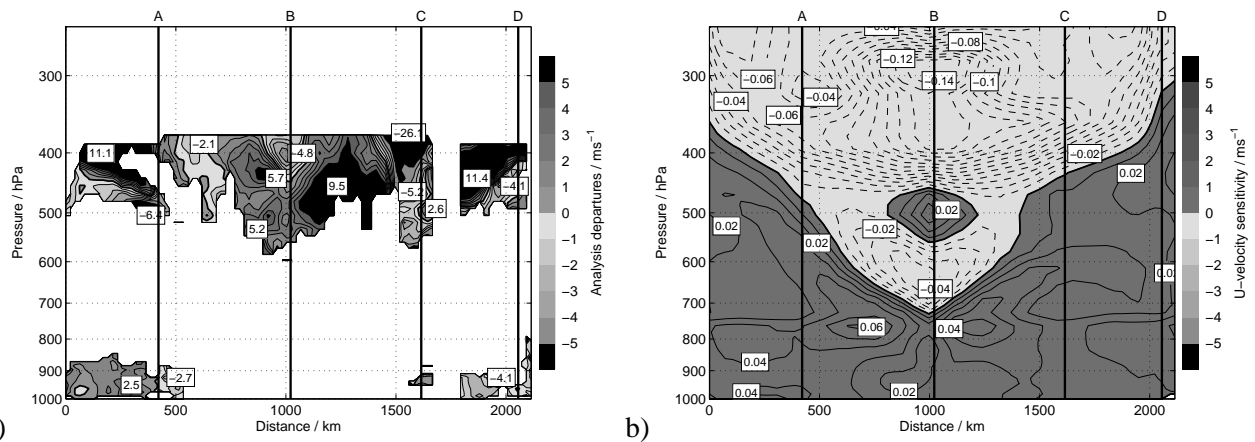


Figure 7: Same as Figs. 1b and 1d, but cross-sections along the flight track of the DLR Falcon on 15 November 2003. Flight track starts and finishes at KEV (see Fig. 6).

therefore, to bring the short-term forecast of a nonlinear model run closer to the true state of the atmosphere by using the information contained in extra observations. For example, CARDINALI et al. (2003) showed that the usage of additional automated profiling aircraft measurements have a substantial positive impact on analysis and forecast accuracy.

A measure for the impact of the DWL observations on the analysis is the computation of analysis differences obtained by calculating the differences of a control run without additional observations and a model run that actively assimilates DWL observations. If accurate additional observations are obtained in regions of large uncertainty, and the assimilation system is properly using the information, the accuracy of the analysis is expected to increase on average. Therefore, analysis differences provide information on AE in terms of a lower bound; further, if it can be assumed that one of the analyses is in a sense better (i.e., more accurate through, e.g., the use of additional accurate observations), then it might be expected that differences are on average pointing into the direction of AEs and indicating the magnitude of AEs.

Fig. 2 shows the u -velocity analysis difference at model level 39 (approximately 500 hPa). The information is spread out over up to 2000 kilometres. The impact varies between -5.1 m/s and 2 m/s with highest amplitudes located downstream. These effects and the periodic sign change of the amplitudes probably result largely from the 4D-Var assimilation scheme, that adjusts model fields to bring the analysis closer to the observations. If a misfit between a previous short-term forecast field and model grid points is observed, the deviation of adjacent points is most likely also too large. Therefore, the assimilation cycle accounts for error correlations within a certain radius around the observations. Comparing the structures of B-norm KAEs (Fig. 3a) and TE-norm KAEs (Fig. 3b), increased sensitivities are found near baroclinic zones. The large equivalent po-

tential temperature gradient with respect to the frontal system at around $45^\circ\text{N}/20^\circ\text{W}$ mainly forms the shape of the sensitivity patterns in this region. In addition, strong flow-dependence is observed (see geopotential height and equivalent potential temperature as shown in Fig. 1a).

However, the B-norm and TE-norm sensitivity patterns are different. The structures of the B-norm KAEs are broader and peak at slightly higher magnitudes (see also, Figs. 1c and 1d). In detail, the minimum amplitude and maximum amplitude of the B-norm sensitivity vary between -1.3 m/s and 0.7 m/s, whereas the TE-norm sensitivity peaks around -0.6 m/s and 0.6 m/s. Closer inspection of the TE-norm sensitivity structures furthermore reveals a more localized organization of minima and maxima. However, the well-pronounced pattern observed in the B-norm sensitivity at around $60^\circ\text{N}/20^\circ\text{W}$ is also visible in the TE-norm pattern, but with lower amplitude. The equivalent potential temperature gradient (see Fig. 1a) is weak in this region, which indicates that these structures presumably arise from instabilities induced by the upper level trough.

The comparison of KAEs and analysis differences (Fig. 2) shows significant discrepancies in terms of spatial structure and magnitude. The analysis differences have more fine-scale spatially organized peaks with higher amplitude. Consequently, the compared structures contain different information about the AE. It can be shown that these differences are of similar character at model levels above and below, and also for the comparison of the v -velocity.

4.3 Statistical comparison

To investigate the hypothesis whether KAEs point towards an improved analysis, correlations between DWL measurements and KAEs were computed. Figs. 4a, 4b, and 4c show scatter plots of analysis departures based

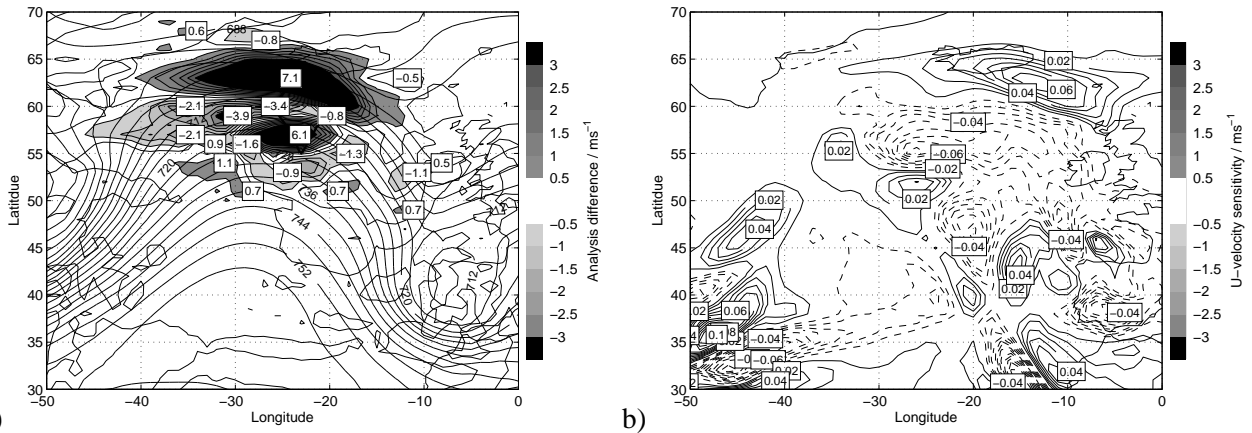


Figure 8: Same as Figs. 2 and 3b, but at model level 37 (approximately 425 hPa). The u -velocity fields are valid at 1800 UTC 15 November 2003.

on passively monitored DWL data and KAEs for a sample of 17226 observations collected on the eight research flights during A-TReC. No significant correlation between analysis departures and KAEs was found (see Figs. 4a and 4b).

In addition, there is only a poor correlation between the B-norm and TE-norm sensitivity perturbations (Fig. 4c), which is an indication for the strong norm-dependence of the adjoint sensitivities. Thus, it is an open question which of these norms – if either of them – is more appropriate to represent AEs. Furthermore, spatial correlation maps with respect to model levels (not shown) confirm the absence of any correlation between analysis differences and KAEs. As pointed out in the introduction, the absence of such correlations might indicate that KAEs are embedded or hidden in the much larger non-growing part of the AE (see, also, GELARO et al. 2000).

To analyse the signal of the analysis difference and the KAEs in more detail, space-averaged (one-dimensional) spectral (in terms of spatial wavenumber) energy densities at model level 39 (approximately 500 hPa) were calculated for the domain $70^{\circ}\text{N}/30^{\circ}\text{W}$ – $50^{\circ}\text{N}/10^{\circ}\text{E}$ valid at 1200 UTC 14 November (Fig. 5a). In this computation, the finite segment of data was detrended by subtracting the linear least-squares regression from the data. The spectral peaks indicate the dominant wavenumbers. It can be seen that the energy spectra of the analysis difference and the TE-norm sensitivity peak at approximately the same wavenumber. Furthermore, their peaks decay more rapidly towards high wavenumbers in contrast to the spectrum of B-norm KAEs. However, the spectral amplitude of the KAEs is much smaller compared than that of the analysis difference.

The vertical distribution of the mean spectral energy density for the same domain and time is shown in Fig. 5b. The spectrum of the TE-norm sensitivity is more uniformly weighted over the model levels, whereas the

spectra of the analysis difference and the B-norm sensitivity show a clear peak at approximately the same model level. This is an implication that the variation of the energy of the analysis difference and the B-norm sensitivity is highest in the region of the atmosphere, which is mainly dominated by the kinetic energy.

In addition, the peak of the analysis difference indicates, that the impact of the DWL measurements is highest around model level 36 (approximately 388 hPa). It can be shown that mean spectral energy densities based on the period 14 November to 28 November 2003 are of same characteristics. Thus, the results shown in Figs. 5a and 5b highlight the observed discrepancies between analysis differences and KAEs in terms of structure and magnitude.

5 Case study for 15 November 2003

5.1 Analysis departures versus KAEs

On 15 November 2003, the DLR Falcon research airplane performed targeted observations, focusing on the dynamically evolving area over the Northern Atlantic to the south of Iceland as shown in Fig. 6. The measurements were collected between 1531 UTC and 1843 UTC. The aim of the targeting event was to reduce the forecast error on a verification domain at verification time 1200 UTC 17 November 2003. The verification domain was located over Southern Europe and Northern Africa ($37^{\circ}\text{N}/02^{\circ}\text{E}$, see Fig. 6). The reason for taking these extra observations was that models predicted heavy rainfall over Southern France, Eastern Spain, and North Africa on 17 and 18 November due to an upper-level low. The low separated from the main trough over the Iberian peninsula on 15 November.

Figs. 7a and 7b show analysis departures based on passively monitored DWL measurements and TE-norm KAEs along the flight track, both for the u -velocity field.

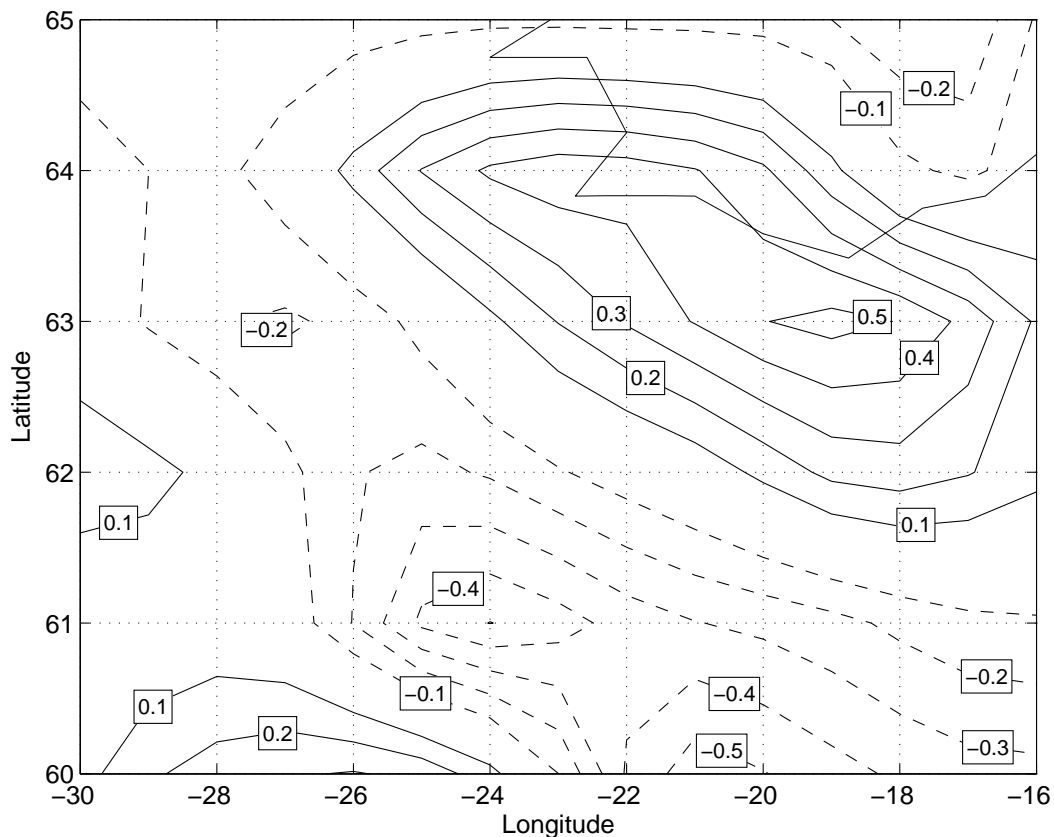


Figure 9: Map of vertical correlations between analysis difference and TE-norm KAEs for the u -velocity field within model level 30 (approximately 200 hPa) and level 45 (approximately 730 hPa) valid at 1800 UTC 15 November 2003. The correlation is computed for the domain $65^{\circ}\text{N}/30^{\circ}\text{W}$ – $60^{\circ}\text{N}/15^{\circ}\text{W}$. Positive values are shown as solid contours and negative values are shown as dashed contours. Contour interval is 0.1.

KAEs are optimized to minimize the forecast error over the predefined verification area as shown in Fig. 6. Similar to Figs. 1c and 1d, no correlation between analysis departures and KAEs (Fig. 7b) is found in terms of spatial structure and magnitude. In detail, the signature of the analysis departures is again characterized by local fine-scale structures with a relative broad amplitude peak spread between -11.4 m/s and 26.1 m/s indicating the source of AEs. In contrast, the TE-norm sensitivity clearly shows more large-scale behaviour. In addition, the sensitivity values vary between -0.11 m/s and 0.16 m/s. These low peaks reflect the influence of iterations on the sensitivity amplitude performed in the minimization process. Similar discrepancies between these quantities also occur for the comparison of the v -velocity (not shown).

5.2 Analysis impact of lidar data

As discussed in section 4.2, the investigation of the impact of extra DWL observations on the analysis may provide useful information about AEs. Fig. 8a shows the u -velocity analysis difference at model level 37 (approximately 425 hPa). Again, information contained in DWL

measurements is spread out over a finite domain surrounding the observation path. The dominant structure at around $60^{\circ}\text{N}/25^{\circ}\text{W}$ ahead of the warm front, which is indicated by closely packed equivalent potential temperature contour lines shown in Fig. 6, is characterized by a meridional amplitude spread of around -3.9 m/s and 7.1 m/s.

Comparing this pattern with the u -velocity TE-norm sensitivity pattern (Fig. 8b), displaced similarities in terms of sign change are apparent. In detail, the prominent, zonal structure with meridional embedded peak-to-peak variations of the analysis difference between 55°N and 65°N at around 25°W also seems to be present in the sensitivity pattern with a much broader area of negative sensitivity values. Furthermore, the positive maximum of the analysis difference at around $63^{\circ}\text{N}/25^{\circ}\text{W}$ is also visible in the sensitivity pattern but shifted eastward by approximately 1000 km.

The observed local minima and maxima of sensitivity values between approximately $30^{\circ}\text{N}/25^{\circ}\text{W}$ and $55^{\circ}\text{N}/25^{\circ}\text{W}$ are clearly linked to baroclinic instabilities induced by the large frontal system located over the Northern Atlantic (see also Fig. 6). Furthermore, a well pronounced sensitivity dipole structure around

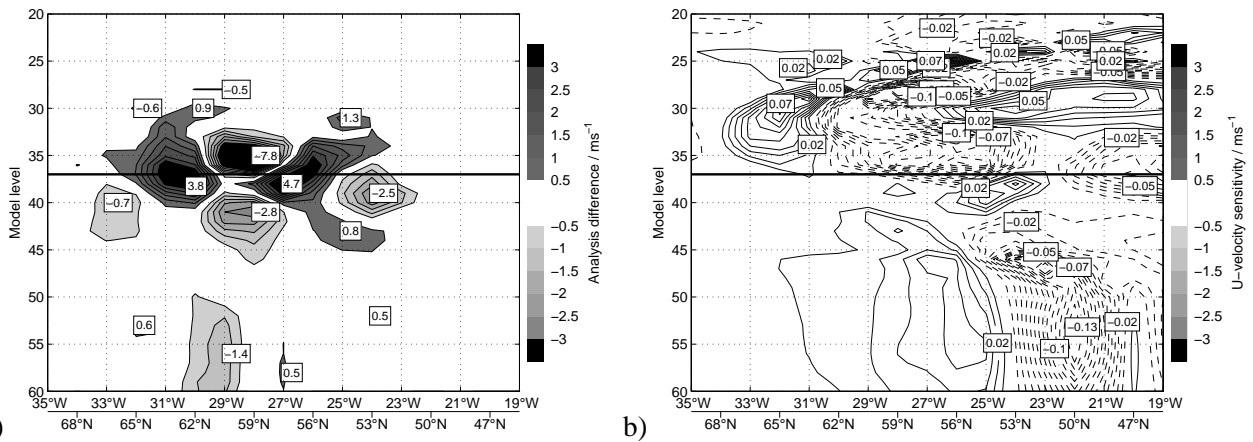


Figure 10: Vertical cross-sections along 70N 35W–45N 19W of analysis difference (ms^{-1}) (a), and TE-norm KAEs (ms^{-1}) (b) for the u -velocity field. Shaded contour interval is 0.5 ms^{-1} . Additional contour interval in panel (b) is 0.01 ms^{-1} , with positive values shown as solid contours and negative values shown as dashed contours. The zero contour is suppressed. Black horizontal line indicates model level 37.

$35^\circ\text{N}/45^\circ\text{W}$ is observed with minimum and maximum peak values of -0.14 m/s and 0.13 m/s . This feature is related to the large equivalent temperature gradient in this area. Further downstream, the temperature gradient is much smaller, which results in lower sensitivity values.

Fig. 9 shows a map of vertical correlations between the analysis difference and KAEs within model level 30 (approximately 200 hPa) and 45 (approximately 730 hPa) for the domain $65^\circ\text{N}/30^\circ\text{W}$ to $60^\circ\text{N}/15^\circ\text{W}$ that contains the observed zonal spread maximum in the horizontal maps (see Figs. 8a and 8b). The low to moderate correlation near $63^\circ\text{N}/19^\circ\text{W}$ indicates that the observed maxima in Figs. 8a and 8b vary at least to some degree in the same direction at different model levels. Thus, one could expect that the structures probably highlight the same source of the AE around this location. However, it remains questionable whether these sensitivity structures represent AEs or whether this correspondence occurs by chance. As a matter of fact, the sensitivity structures show a closer resemblance to the analysis difference than has been the case for 1200 UTC 14 November (Figs. 2, 3a and 3b).

Figs. 10a and 10b show NW-SE cross-sections of the u -velocity analysis difference and KAEs, respectively. The largest impact is found between model level 30 (approximately 200 hPa) and model level 45 (approximately 728 hPa) with minimum and maximum amplitude peaks around -7.8 m/s and 4.7 m/s . The investigation of further cross-sections indicates that the impact of DWL measurements varies strongly in the vertical, which is mainly due to rapidly changing properties of the atmosphere (e.g., temperature) in the vertical, resulting in a correlation radius that also strongly varies with height. The corresponding cross-section of the sensitivity pattern (Fig. 10b) shows backward tilted bands at lower model levels, which are a characteristic feature of

KAEs (ISAKSEN et al., 2005). As in the comparison of analysis departures and KAEs, the comparison of analysis differences and TE-norm KAE patterns reveals significant differences.

6 Summary and discussion

In this study, substantial differences are documented in terms of spatial structure and magnitude of analysis departures based on passively monitored DWL data and KAEs based on different initial time norms. Similar discrepancies are observed between analysis differences and KAEs. The u - and v -velocity KAE sensitivity values are much smaller and the locations of their maxima and minima are different, when compared to those of the analysis departures and analysis differences.

In addition, neither quantitative, nor qualitative correlation was found. It turned out, that the squared correlation coefficient between analysis departures and KAEs is overall zero for all cases with lidar observations. Moreover, the correlation between analysis differences and KAEs in physical space is poor. It was further found that small-scale structures are absent in KAE sensitivity fields which presumably is to some degree related to the reduced resolution and the use of simplified physics in the algorithm used for the computation of the KAEs.

It might be attempted to interpret these results by maintaining that KAEs carry no information about "true" AEs and are an artefact of the methods used to compute them. This interpretation is more in line with the conclusions offered by ISAKSEN et al. (2005) and would in fact amount to rejecting the hypothesis stated in the introduction.

However, a careful consideration of the results presented here indicates that they provide no evidence for rejecting the initial hypothesis that the KAEs are embedded in the AEs and thus are a small, but important

part of AE. In fact, the substantial differences in magnitude and structure – as documented here – between KAEs and the approximations of AE used here do not necessarily contradict this hypothesis. Previous studies carried out in the context of state-of-the-art assimilation systems explicitly mention such differences in magnitude and structure (e.g., GELARO et al. 2000). The fact that KAEs – when time-evolved – do reduce forecast error, as well as present knowledge about nonmodal finite-time error-growth evolution are also both consistent with this hypothesis.

An additional caveat and final comment needs to be mentioned. The weak correlation found between B-norm and TE-norm KAEs does leave the question open as to what kinds of constraints are in fact suitable to define KAEs. Nevertheless, even in view of this uncertainty, it is evident that KAEs represent some unstable components of the analysis error in phase space that are linked to regions of large instability in physical space. It is quite clear though that further studies are strongly needed to elucidate the precise physical nature of KAEs.

Acknowledgements

The authors are very grateful to Carla CARDINALI (ECMWF) for the set-up of several experiments and to Thomas JUNG (ECMWF) for computing KAEs. Furthermore, we would like to acknowledge all the people at DLR responsible for obtaining these unique DWL observations during A-TReC, in particular Andreas DÖRNBRACK, Stephan RAHM, and Rudolf SIMMET. The comments by four reviewers on earlier versions of this paper have been very constructive and helpful for the authors' attempts to improve the description of this work. In addition, the first author is grateful for the having had the opportunity of a long-term visit at the Institute of Atmospheric Physics, DLR, during Summer 2006 and for the financial support provided by the International Relations Office of the University of Innsbruck.

References

- BUZZA, R., 1994: Localization of optimal perturbations using a projection operator. – *Quart. J. Roy. Meteor. Soc.* **120**, 1647–1682.
- CARDINALI, C., L. ISAKSEN, E. ANDERSSON, 2003: Use and Impact of Automated Aircraft Data in a Global 4DVAR Data Assimilation System. – *Mon. Wea. Rev.* **131**, 1865–1877.
- CARON, J.-F., M.K. YAU, S. LAROCHE, P. ZWACK, 2007: The Characteristics of Key Analysis Errors. Part I: Dynamical Balance and Comparison with Observations. – *Mon. Wea. Rev.* **135**, 249–266.
- COURTIER, P., J.N. THEPAUT, A. HOLLINGSWORTH, 1994: A strategy for operational implementation of 4D-VAR, using an incremental approach. – *Quart. J. Roy. Meteor. Soc.* **120**, 1367–1388.
- COURTIER, P., E. ANDERSSON, W. HECKLEY, D. VASILJEVIC, M. HAMRUD, A. HOLLINGSWORTH, F. RABIER, M. FISHER, J. PAILLEUX, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation – *Quart. J. Roy. Meteor. Soc.* **124**, 1783–1807.
- EHRENDORFER, M., J.J. TRIBBIA, 1997: Optimal Prediction of Forecast Error Covariances through Singular Vectors. – *J. Atmos. Sci.* **54**, 286–313.
- ERRICO, R.M., 1997: What Is an Adjoint Model? – *Bull. Amer. Meteor. Soc.* **78**, 2577–2591.
- , 2000: Interpretations of the total energy and rotational energy norms applied to determinations of singular vectors. – *Quart. J. Roy. Meteor. Soc.* **126**, 1581–1599.
- ERRICO, R.M., R. YANG, M. MASUTANI, J.S. WOOLLEN, 2007: The use of an OSSE to estimate characteristics of analysis error. – *Meteorol. Z.* **16**, 695–708.
- GELARO, R., C.A. REYNOLDS, R.H. LANGLAND, G.D. ROHALY, 2000: A Predictability Study Using Geostationary Satellite Wind Observations during NORPEX. – *Mon. Wea. Rev.* **128**, 3789–3807.
- GELARO, R., C.A. REYNOLDS, R.M. ERRICO, 2002: Transient and asymptotic perturbation growth in a simple model. – *Quart. J. Roy. Meteor. Soc.* **128**, 205–227.
- JUNG, T., J. BARKMEIJER, 2006: Sensitivity of the Tropospheric Circulation to Changes in the Strength of the Stratospheric Polar Vortex. – *Mon. Wea. Rev.* **134**, 2191–2207.
- ISAKSEN, L., M. FISHER, E. ANDERSSON, J. BARKMEIJER, 2005: The structure and realism of sensitivity perturbations and their interpretation as “Key Analysis Errors”. – *Quart. J. Roy. Meteor. Soc.* **131**, 3053–3078.
- KLINKER, E., F. RABIER, R. GELARO, 1998: Estimation of key analysis errors using the adjoint technique. – *Quart. J. Roy. Meteor. Soc.* **124**, 1909–1933.
- LEUTBECHER, M., 2003: Adaptive observations, the Hessian metric and singular vectors. – *Proc. ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean*, Reading, UK, 8–12 September 2003.
- MARSEILLE, G.J., A. STOFFELEN, J. BARKMEIJER, submitted: Sensitivity Observing System Experiment (SOSE) – A New Effective NWP-based Tool in Designing the Global Observing System. – *Tellus A*.
- RABIER, F., E. KLINKER, E. COURTIER, P., A. HOLLINGSWORTH, 1996: Sensitivity of forecast errors to initial conditions. – *Quart. J. Roy. Meteor. Soc.* **122**, 121–150.
- RABIER, F., H. JÄRVINEN, E. KLINKER, J.F. MAHOUF, A. SIMMONS, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part 1: experimental results with simplified physics. – *Quart. J. Roy. Meteor. Soc.* **126**, 1143–1170.
- WEISSMANN, M., C. CARDINALI, 2007: Impact of airborne Doppler lidar observations on ECMWF forecasts. – *Quart. J. Roy. Meteor. Soc.* **133**, 107–116.
- WEISSMANN, M., R. BUSEN, A. DÖRNBRACK, S. RAHM, O. REITEBUCH, 2005: Targeted Observations with an Airborne Wind Lidar. – *J. Atmos. Oceanic Technol.* **22**, 1706–1719.